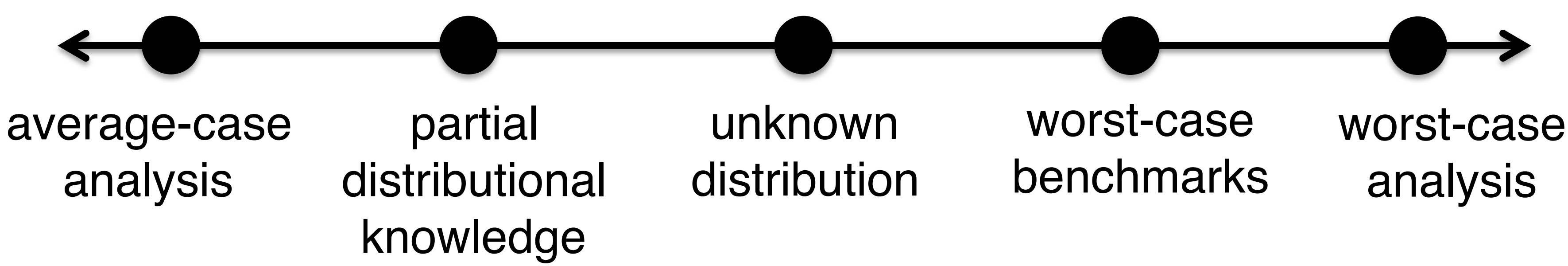
Outposts Between Worst- and Average-Case Analysis

a case study in auction design

Tim Roughgarden (Stanford University)



Worst-Case Analysis

Worst-case analysis: $cost(A) := \sup_z cost(A,z)$

– cost(A,z) = performance of algorithm A on input z

Pros of WCA: relatively analytically tractable

- universal applicability (no input assumptions)
- countless killer applications

Cons of worst-case analysis: overly pessimistic

- can rank algorithms inaccurately (LP, paging)
- no data model (rather: “Murphy’s Law” model)

Average-Case Analysis

Average-case analysis: $cost(A) := E_z\ cost(A,z)$

– for some distribution over inputs z

- well motivated if:
  - (i) detailed and stable understanding of distribution;
  - and (ii) don't need a general-purpose solution

Concern: advocates brittle solutions overly tailored to input distribution.

– which might be wrong, change over time, or be different in different applications

Hybrid Models

Thesis: for many problems there is a "sweet spot" between worst- and average-case analysis.

– where unknown distribution D lies in some known set

worst-case

$\sup_x \text{cost}(A,x)$

average-case

$E_x[\text{cost}(A,x)]$

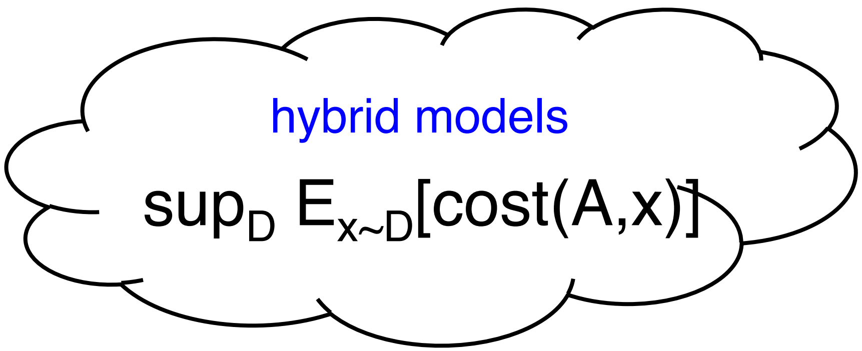
hybrid models

$\sup_D E_{x\sim D}[\text{cost}(A,x)]$

Hybrid Models

Thesis: for many problems there is a "sweet spot" between worst- and average-case analysis.

– where unknown distribution D lies in some known set



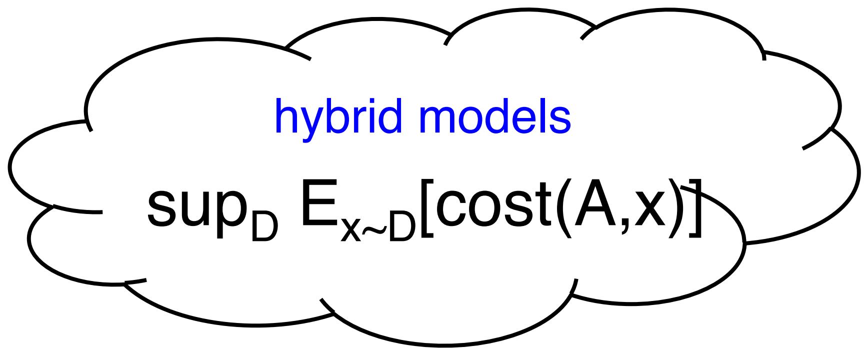
Benefits:

- robust near-optimality guarantees
- via natural algorithms

Hybrid Models

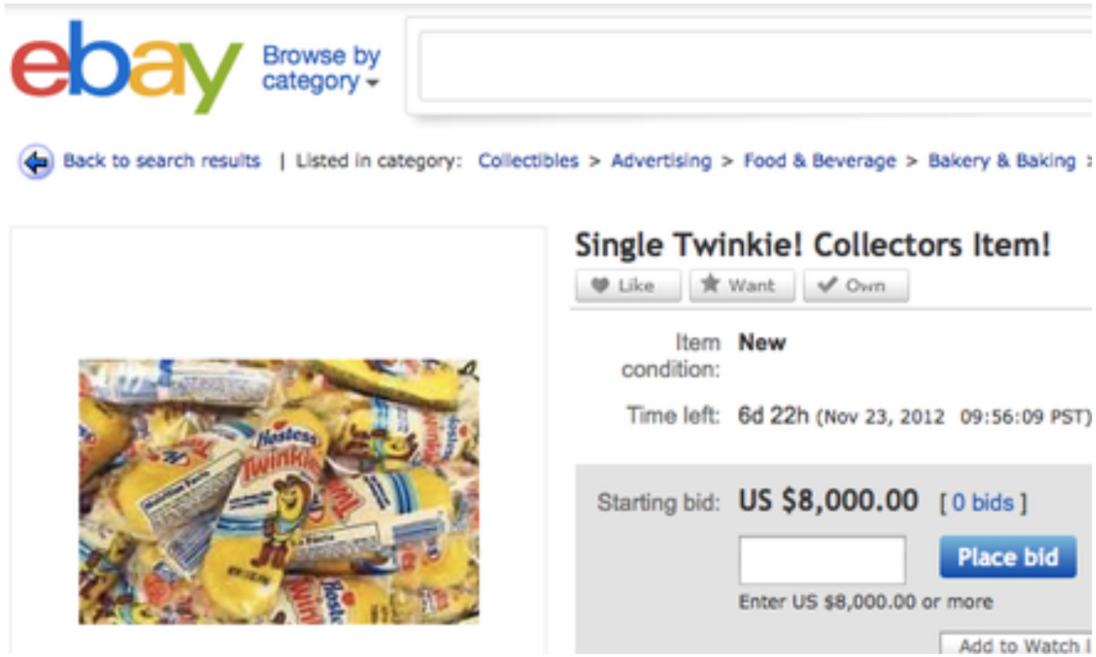
Thesis: for many problems there is a “sweet spot” between worst- and average-case analysis.

– where unknown distribution D lies in some known set



Examples: smoothed analysis, semi-random models, random-order models (secretary), etc.

Case Study: Single-Item Auctions



Question: what opening bid maximizes revenue?

Issue: depends on what bidders are willing to pay (which is unknown).

Worst-Case Analysis Fails

Single-item auction setup:

- 1 seller with 1 item; n bidders
- bidder i has private valuation v_i
- goal: maximize revenue

Example: suppose one bidder with valuation v.

- optimal opening bid = v (=> revenue = v)
- no opening bid guaranteed to obtain near-optimal revenue for all possible values of v

Bayesian Approach (Economics)

Bayesian assumption: bidders' valuations $v_1,...,v_n$ drawn independently from distributions $F_1,...,F_n$.

- F_i's known to seller, v_i's unknown
- need minor conditions on F_i's (won't discuss)

Goal: find auction with most expected revenue.

- collect one bid per bidder
- decide on a winner (if any) and a selling price
- ex: eBay = second-price auction with a reserve

Optimal Single-Item Auctions

[Myerson 81]: characterized the optimal auction, as a function of the prior distributions $F_1,\ldots,F_n$.

I.i.d. case: ($F_1=F_2=\ldots=F_n$) optimal auction = 2nd-price auction with suitable reserve price.

- reserve = *monopoly price* ($\mathrm{argmax}_p\, p(1-F(p))$)

General case: complicated, depends in detailed way on $F_1,\ldots,F_n$.

Perils of Average-Case Analysis

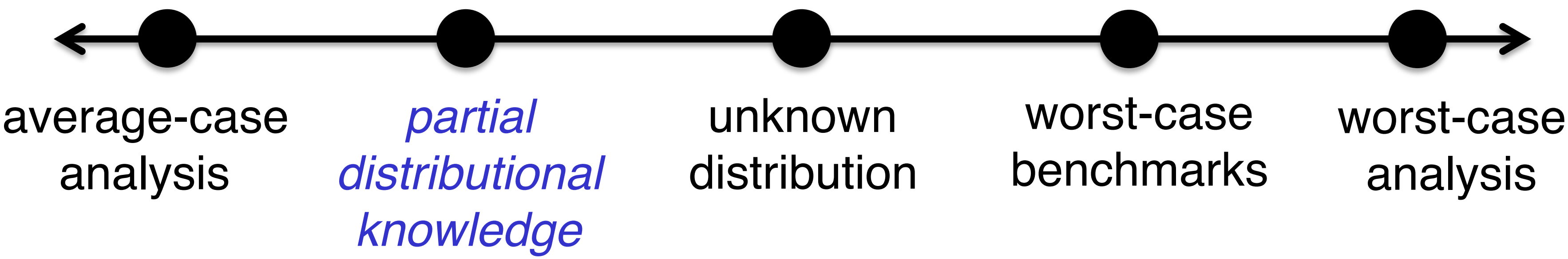
Recall concern: average-case analysis might advocate brittle solutions specific to distribution.

Wilson's doctrine: auction theory should advocate "robust" solutions that do not depend on the details of the knowledge assumptions.

"Only by repeated weakening of common knowledge assumptions will the theory approximate reality."

(Robert Wilson, 1987)

Road Map


average-case analysis
partial distributional knowledge
unknown distribution
worst-case benchmarks
worst-case analysis

Models of Partial Knowledge

Idea: design near-optimal auctions that require only "minimal" distributional knowledge.

- simple statistics (median, etc.), or samples

Interpretations:

1. Don't have detailed distributional knowledge.
2. Want simple and robust auctions that don't exploit detailed distributional knowledge.

Reserve Price-Based Auctions

Interpretations of [Hartline/Roughgarden 09]:

1. Want to use reserve price-based auctions.
2. Only know monopoly prices of distributions.
   - I.i.d. case => already enough to implement opt
   - non-i.i.d. case => not enough to implement opt

Goal: reserve price-based auction with expected revenue close to optimal (no matter what the prior $F_1,\ldots,F_n$ is).

Simple vs. Optimal Auctions

Theorem: [Chawla/Hartline/Kleinberg 07] [Hartline/Roughgarden 09] second-price auction with (bidder-specific) monopoly reserve prices has expected revenue at least 50% of optimal.

- best-possible guarantee for reserve-price-based auctions (for worst-case prior $F_1,\ldots,F_n$)

Techniques: uses analytical tools from traditional optimal auction theory. ("virtual values")

Statistics as Partial Knowledge

- [Alaei/Hartline/Niazadeh/Pountourakis/Yuan 15] guarantee of 36% for *anonymous* reserve price (eBay)
- [Azar/Micali 12] assume knowledge only of mean and variance of distributions
- [Azar/Daskalakis/Micali/Weinberg 13] assume knowledge only of median of distributions
- many more papers on more complex problems

Samples as Partial Knowledge

A PAC Model: [Cole/Roughgarden 14]

- inspired by learning theory [Vapnik/Chervonenkis 71, Valiant 84]
- given i.i.d. samples from unknown distributions $F_1,...,F_n$ (e.g., past bids)
- goal: $(1-\varepsilon)$-approximation of optimal revenue

Question: How much data (# of samples) is necessary and sufficient?

- relevant to practice [Ostrovsky/Schwarz 09]

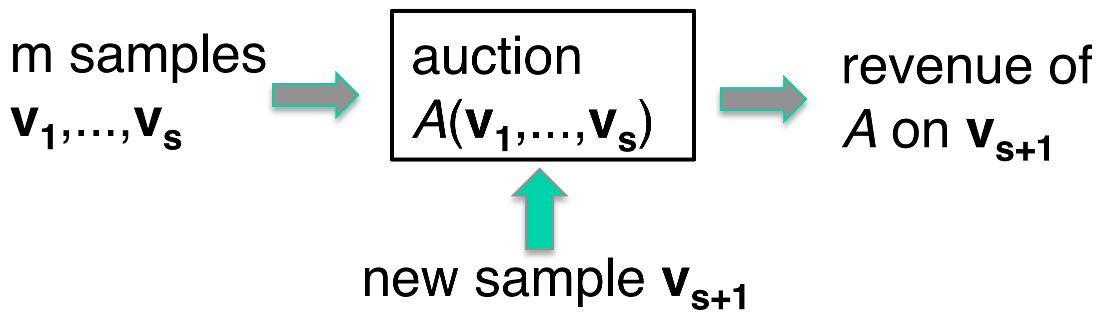
Formalism

Step 1: seller gets s samples $\mathbf{v}_1,\ldots,\mathbf{v}_s$ from $\boldsymbol{F}$

– each $\mathbf{v}_i$ an *n*-vector (one valuation per bidder)

Step 2: seller picks single-item auction $A = A(\mathbf{v}_1,\ldots,\mathbf{v}_s)$

Step 3: A run on a fresh sample $\mathbf{v}_{s+1}$ from $\boldsymbol{F}$



Goal: design A so $E_{v_1,\ldots,v_s}[E_{v_{s+1}}[\mathrm{Rev}(A(v_1,\ldots,v_s)(v_{s+1}))]]$ close to OPT (for $\boldsymbol{F}$).

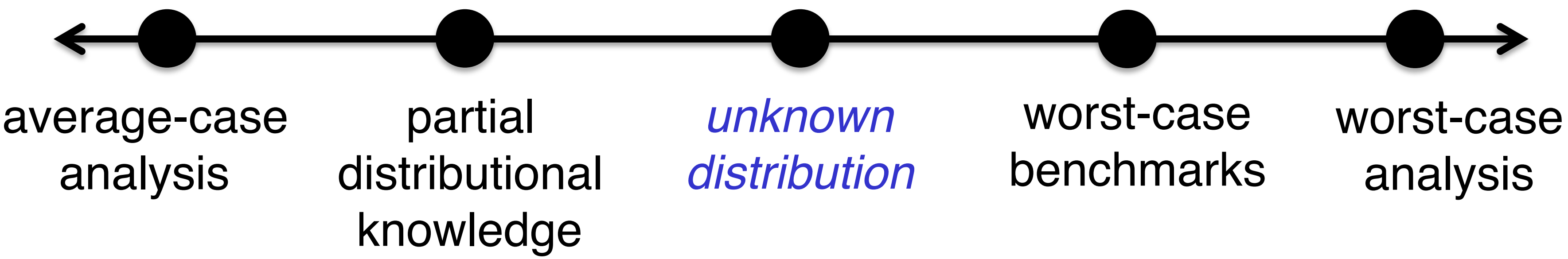
Representative Results

Good news: [Huang/Manour/Roughgarden 15]
I.i.d. case: Polynomial (in ε^{-1} only) samples suffice for a (1-ε)-approximation.

Mixed news: [Cole/Roughgarden 14], [Morgenstern/Roughgarden 15], [Devanur/Huang/Psamos 16]
Non-i.i.d. case: poly(n,ε^{-1}) samples necessary and sufficient for a (1-ε)-approximation.

Corollary (of proof): near-optimal auctions require detailed distributional knowledge.

Road Map



average-case analysis
partial distributional knowledge
unknown distribution
worst-case benchmarks
worst-case analysis

Bulow-Klemperer Theorem

Setup: single-item auction, valuations i.i.d. from F.

Theorem [Bulow-Klemperer 96]: for every $n \geq 1$:

Vickrey's revenue	$\geq$	OPT_F's revenue
[with (n+1) i.i.d. bidders]		[with n i.i.d. bidders]

Interpretation: small increase in competition more important than running optimal auction.

Prior-Independent Auctions

Goal: [Dhangwotnotai/Roughgarden/Yan 10] *prior-independent auction* = expected revenue almost as good as if distribution was known up front

- no matter what the distribution is
- distribution used only in the *analysis* of the auction, not in its *design*

hybrid models

$\sup_D E_{x\sim D}[\text{cost}(A,x)]$

The Setup

- n bidders, known partition into groups $G_1,\ldots,G_t$
  - let $k = \min_h |G_h|$; provably need $k \geq 2$

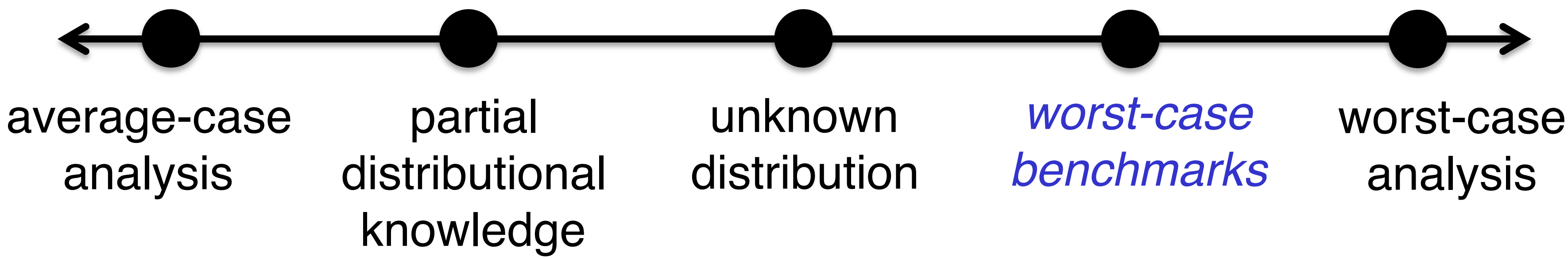
- valuations of bidders in G_h drawn IID from *unknown* distribution F_h

The Single Sample Mechanism

1. for each group G_h, pick a *reserve bidder* i_h uniformly at random
2. let i = highest non-reserve bidder
3. sell item to i if and only if her bid is at least that of the reserve bidder from her group

Theorem: [Dhangwotnotai/Roughgarden/Yan 10] expected revenue at least a $\frac{1}{2} \circ (k-1)/k$ fraction of optimal (recall $k = \min_h |G_h|$)

Road Map


average-case analysis
partial distributional knowledge
unknown distribution
worst-case benchmarks
worst-case analysis

Toward Worst-Case Analysis

Goal: *prior-free* auction. (input-by-input guarantees)

"Desired Theorem: for every valuation profile v: auction A's revenue on v is at least OPT(v)/α."

(for a hopefully small constant α)

Question: how to define OPT(v)?

– recall failure of worst-case analysis

Idea: [Goldberg/Hartline/Wright 01] define suitable *revenue benchmarks.*

Template for Benchmarks

Result: [Hartline/Roughgarden 08] general template for meaningful worst-case auction benchmarks.

Average-case thought experiment: suppose every valuation drawn i.i.d. from a distribution F.

- optimal = 2^{nd}-price auction with suitable reserve

Benchmark: define OPT(v) = max revenue obtained on v by auction that is optimal for some F.

- deterministic proxy for i.i.d. bidders
- analogs: no-regret learning, static optimality

Template for Benchmarks

Benchmark: define OPT(v) = max revenue obtained by an auction that is optimal for some F.

- regenerates previously studied benchmarks [Goldberg/Hartline/Wright 01]
- gives new benchmarks for more general problems [Hartline/Roughgarden 08], [Devanur/Hartline 09], [Hartline/Yan 11], [Leonardi/Roughgarden 12]

Meaning: approximate benchmark input-by-input => approximate all Bayesian optimal auctions. (i.e., prior-free => prior-independent)

Summary

- worst-case analysis: can be overly pessimistic, give no advice about which auction to run
- average-case analysis: can be overly brittle, advocate overly complex solutions
- outposts on the road in between:
  - assume knowledge only of simple statistics (monopoly price, median, mean + variance)
  - assume only a polynomial number of samples
  - prior-independent auctions
  - prior-free auctions w.r.t. revenue benchmarks